§**5.4.4. An SMLE approach**

Recall we consider regression model $X_i = \beta \mathbf{z}_i + W_i$, $i = 1, ..., n$, where $W_i = \alpha + \epsilon_i$. Yu and Wong (2003a,b,c) proposed the SMLE of (β, S_o), based on complete data, RC data and IC data, respectively. The semi-parametric likelihood function is

$$Ł = \prod_{i=1}^{n} \mu_F(I_i - \mathbf{b}\mathbf{z}_i), \tag{4.1}$$

where F is a cdf, I_i's are the observed intervals containing X_i, and $I_i - c$ is a shift of I_i by c units. The SMLE of (β, F_o) is

$$(\hat{\beta}, \hat{F}) = \text{argmax}_{\mathbf{b},F}\{Ł(\mathbf{b}, F) : \ \mathbf{b} \in \mathcal{R}^p, \ F(t) \uparrow, \ F(-\infty) = 0, F(\infty) = 1\}.$$

Remark. With complete data, the LSE **minimizes**

$$\sum_{i=1}^{n}(X_i - \alpha - \mathbf{b}z_i)^2 \text{ or } \int t^2 d\hat{F}_\epsilon(t).$$

The SMLE of (α, β, F_o) **maximizes**

$$Ł(a, \mathbf{b}, F) = \prod_{i=1}^{n} f_\epsilon(X_i - a - \mathbf{b}z_i).$$

Example 1. Consider a simple example of RC data, say,

3 (M_i, δ_i, z_i)s are $(1, 1, 1)$, $(2, 0, 1)$, $(3, 1, 0)$.

$$\begin{aligned}
\text{Let } T_i(\beta) = M_i - \beta z_i \qquad & (X_i - \beta Z_i \sim F_o(t) = F_{X|Z}(t|0).\\
I_i\text{'s are } \{X_1\}, (X_2, \infty), \{X_3\} & \\
I_i - bz_i\text{'s are } \{T_1(b)\}, (T_2(b), \infty), \{T_3(b)\}. & \\
Ł(b, F) = \prod_{i=1}^{n} \mu_F(I_i - bz_i) = f(1-b)S(2-b)f(3) \qquad & \textbf{Why ?}\\
Ł(b, F) \le \hat{f}_b(1-b)\hat{S}_b(2-b)\hat{f}_b(3) &
\end{aligned}$$

where $\hat{S}_b(t)$ is the PLE based on $T_1(b)$, $T_2(b)$, $T_3(b)$, and $\hat{f}_b(t)$= ?

The PLE $\hat{S}_b$ depends on the ranks of $T_i(b) = M_i - bz_i$'s,

Let (r_1, r_2, r_3) be the ranks of $\begin{matrix}(T_1 & T_2+ & T_3)(b)\\ (\ 1-b & (2-b)+ & 3\ )\end{matrix}$

$T_i(b)$'s change their ranks after their ties: $M_k - bz_k = M_j - bz_j$ with $z_k \neq z_j$.

$1 - b = 3$ $(T_1(b) = T_3(b))$ and $2 - b = 3$ $(T_2(b) = T_3(b))$.

Their solutions are $b_1 = -2$ and $b_2 = -1$.

-2 and -1 partition $(-\infty, \infty)$ into 5 disjoint intervals.

r_1, r_2, r_3 remain constant in each interval.

$b \in$	$(-\infty, -2)$	$\{-2\}$	$(-2, -1)$	$\{-1\}$	$(-1, \infty)$
$r_i s$	$(2, 3, 1)$	$(1.5, 3, 1.5)$	$(1, 3, 2)$	$(1, 3, 2)$	$(1, 2, 3)$
$\mu_{\hat{F}_b}(I_i - bz_i)s$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{2}{3}, \frac{1}{3}, \frac{2}{3})$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$
$\max Ł$	*no*	*yes*	*no*	*no*	*yes*

Each point in $\{-2\} \cup (-1, \infty)$ is an SMLE of β. For examples

$b \in$	$(-\infty, -2)$	$\{-2\}$	$(-2, -1)$	$\{-1\}$	$(-1, \infty)$
$b =$	-3	-2	-1.5	-1	0
$\vec{T}(b) = (1-b, (2-b)+, 3)$	$(4, 5+, 3)$	$(3, 4+, 3)$	$(2.5, 3.5+, 3)$	$(2, 4+, 3)$	$(1, 2+, 3)$
$r_i s$	$(2, 3, 1)$	$(1.5, 3, 1.5)$	$(1, 3, 2)$	$(1, 3, 2)$	$(1, 2, 3)$
$\mu_{\hat{F}_b}(I_i - bz_i)s$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{2}{3}, \frac{1}{3}, \frac{2}{3})$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$
$\max Ł$	*no*	*yes*	*no*	*no*	*yes*

One SMLE of (β, F_o) is $(\hat{\beta}, \hat{F}_{\hat{\beta}}(t)) = (-2, \frac{2}{3}\mathbf{1}(t \geq 3))$.
How about $(-2, \frac{2}{3}\mathbf{1}(t \geq 3)) + \frac{1}{3}\mathbf{1}(t \geq 5)$? $(-2, \frac{2}{3}\mathbf{1}(t \geq 3)) + \frac{1}{3}\mathbf{1}(t \geq 4)$?

Example 2. Consider the simple linear regression ($p = 1$), with complete data, say $X = \beta Z + W$, where W and $Z \sim Bin(1, 0.5)$ and $\beta = 1$. The possible values of the observation (Z, X) are $(0, 0)$, $(0, 1)$, $(1, 1)$ and $(1, 2)$, denoted by (Z_i, X_i), $i = 1$, ..., 4. Thus there are 4 possible values of $T(b)$, say, $T_i(b) = X_i - bZ_i$: 0, 1, $1-b$, $2-b$. Suppose a random sample of size n contains N_1, N_2, N_3 and N_4 of them. One may consider

the parametric approach (say the MLE or the MME of (p, β), assuming $W \sim Bin(1, p)$;
the semi-parametric approach (the LSE of β, assuming F_W is unknown);
the SMLE approach (assuming F_W is unknown) as follows.

The empirical df

$$\hat{f}_b(T_i(b)) = \begin{cases} \frac{N_i}{n} & \text{if } T_i(b) \neq T_j(b) \ \forall\ j \neq i, \\ \frac{N_i + N_j}{n} & \text{if } T_i(b) = T_j(b) \text{ for only one } j \neq i \\ ? & \text{otherwise} \end{cases} \text{ where } i, j = 1, 2, 3, 4. \tag{4.2}$$

The possible solutions b to the equations $T_i(b) = T_j(b)$ are 0, 1, 1 and 2.

They partition $(-\infty, \infty)$ into 7 intervals.

$$\begin{pmatrix} b: & 0 & 1 & 2 & OW \\ (T_1, ..., T_4): & (0, 1, 1-b, 2-b) = (0, 1, 1, 2) & 0, 1, 0, 1 & 0, 1, -1, 0 & \textit{no ties} \\ \hat{f}_{T(b)}: & (\frac{N_1}{n})^{N_1}, (\frac{N_2+N_3}{n})^{N_2}, (\frac{N_2+N_3}{n})^{N_3}, (\frac{N_4}{n})^{N_4} & & & \end{pmatrix}$$

For n large enough, $N_i \approx n/4$, and likelihood function (1.2) or (4.1)
$\mathcal{L} = \prod_{i=1}^n \hat{f}_b(T_i(b)) =$

$$\begin{cases} (\frac{N_1+N_3}{n})^{N_1+N_3} (\frac{N_2+N_4}{n})^{N_2+N_4} \approx (0.5)^n & \text{if } b = 1, \\ (\frac{N_1}{n})^{N_1} (\frac{N_2+N_3}{n})^{N_2+N_3} (\frac{N_4}{n})^{N_4} \approx (0.5)^{n/2} (0.25)^{n/2} & \text{if } b = 0, \\ (\frac{N_1+N_4}{n})^{N_1+N_4} (\frac{N_2}{n})^{N_2} (\frac{N_3}{n})^{N_3} \approx (0.5)^{n/2} (0.25)^{n/2} & \text{if } b = 2, \\ (\frac{N_1}{n})^{N_1} (\frac{N_2}{n})^{N_2} (\frac{N_3}{n})^{N_3} (\frac{N_4}{n})^{N_4} \approx (0.25)^n & \text{otherwise,} \end{cases} \tag{4.3}$$

is maximized by $b = 1$. Thus the SMLE of β is $\hat{\beta} = 1$ for all large n. consistent ? efficient ?
The SMLE of F_o is $\hat{F}(t) = \hat{p}\mathbf{1}(t \geq 0) + (1 - \hat{p})\mathbf{1}(t \geq 1)$, where $\hat{p} = (N_1 + N_3)/n$.
The LSE $\tilde{\beta}_{LSE} = \beta + \frac{N_1 N_4 - N_2 N_3}{(N_1+N_4)(N_2+N_3)}$. Thus,

$\mathrm{P}(\tilde{\beta}_{LSE} \neq \beta \text{ i.o.}) = 1$.

The LSE $\tilde{\beta}_{LSE}$ satisfies

$\sqrt{n}(\tilde{\beta}_{LSE} - \beta) \xrightarrow{D} N(0,1)$, and $n\sigma^2_{\tilde{\beta}_{LSE}} \to 1$ as $n \to \infty$,
as the asymptotic variance $\sigma^2_{\tilde{\beta}_{LSE}} = \frac{1}{n}$.

It can be shown that $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} N(0,0)$ and $n\sigma^2_{\hat{\beta}} \to 0$.
The MLE does not have this property.□

Remark: If $p > 1$, the solution to SMLE is little bit more complicated, but doable.